

AI Based CO-PO Mapping using NLP and ML

Yash Guram

Department of Computer Engineering
Watumull Institute of Engineering and
Technology Thane, India

Sakshi Gupta

Department of Computer Engineering
Watumull Institute of Engineering and
Technology Thane, India

Darshan Kale

Department of Computer Engineering
Watumull Institute of Engineering and
Technology Thane, India

Rushiraj Kadam

Department of Computer Engineering
Watumull Institute of Engineering and
Technology Thane, India

Prof. Sandeep More
Computer Engineering
Watumull Institute of Engineering and
Technology Thane, India

Abstract

The accurate mapping of Course Outcomes (COs) to Program Outcomes (POs) is a critical, yet labour-intensive and subjective process in Outcome-Based Education (OBE) frameworks mandated by accreditation bodies like NBAnd ABET. Traditional manual mapping by faculty is prone to inconsistencies, while early automated systems relying on keyword matching (TF-IDF) or static word embeddings (Word2Vec) fail to capture nuanced semantic relationships. This paper proposes a novel automated framework that leverages the Bidirectional Encoder Representations from Transformers (BERT) model to perform semantic similarity analysis between CO and PO statements. By generating contextual embeddings and then feeding them into an eXtreme Gradient Boosting (XGBoost) classifier, the framework accurately constructs a CO-PO correlation matrix. The system further integrates this mapping with a structured attainment calculation model, automating the entire OBE assessment cycle. Experimental validation demonstrates that our BERT-based approach, enhanced by XGBoost, significantly outperforms traditional NLP techniques, achieving higher accuracy and consistency. The proposed system reduces faculty workload, ensures objectivity, and provides a reliable, data-driven tool for continuous quality improvement in engineering education.

Keywords—Outcome-Based Education (OBE), CO-PO Mapping, Natural Language Processing (NLP), BERT, Attainment Analysis, Educational Data Mining, Accreditation, XGBoost.

I. Introduction

Outcome-Based Education (OBE) has become the cornerstone of modern engineering education, shifting

the focus from traditional input-centric models to a structured framework emphasizing demonstrable outcomes [1]. Within this paradigm, **Course Outcomes (COs)** define specific competencies students should acquire from individual courses, while **Program Outcomes (POs)** represent the broader, overarching attributes graduates are expected to possess [2]. Establishing a precise correlation between COs and POs is not merely an administrative task but a fundamental requirement for evaluating curriculum effectiveness and ensuring alignment with accreditation standards such as those set by the National Board of Accreditation (NBA) and the Accreditation Board for Engineering and Technology (ABET).

The conventional methodology for CO-PO mapping relies heavily on **manual execution by faculty members**. This process, while leveraging domain expertise, introduces significant challenges including subjectivity, inconsistency across evaluators, and a substantial expenditure of time and effort [3]. As institutions grow and curricula evolve, the scalability of manual mapping becomes a severe bottleneck. These limitations have spurred the exploration of computational solutions, particularly from the field of Natural Language Processing (NLP).

Early automated attempts utilized techniques like **Term Frequency-Inverse Document Frequency (TF-IDF)** for keyword-based matching and **Word2Vec** for semantic similarity [4]. However, these models operate on a superficial level. TF-IDF ignores semantic meaning, treating synonyms as distinct entities, while Word2Vec generates context-independent word embeddings, failing to grasp the intent behind longer, more complex

educational statements [5]. The recent advent of **transformer-based models**, especially BERT (Bidirectional Encoder Representations from Transformers), has revolutionized NLP by providing deep, contextual understanding of language [6].

This paper presents a comprehensive framework that harnesses the power of **BERT** combined with an **eXtreme Gradient Boosting (XGBoost)** model to automate and enhance the CO-PO mapping and attainment analysis process. Our core contributions are:

1. The development of an automated mapping pipeline using **BERT embeddings** to capture deep semantic relationships, followed by an **XGBoost classifier** to generate the final CO-PO correlation matrix with high accuracy.
2. The integration of this mapping with a robust attainment analysis module, creating an end-to-end OBE evaluation system.
3. comparative analysis demonstrating the superiority of our combined BERT-based approach over traditional NLP methods like TF-IDF, Word2Vec, and SpaCy.

The rest of the paper is organized as follows: Section II reviews related work. Section III details the proposed framework architecture, focusing on the BERT and XGBoost integration. Section IV discusses the methodology for attainment analysis. Section V presents a comparative analysis of results. Finally, Section VI concludes the paper and outlines future work.

II. Literature Survey

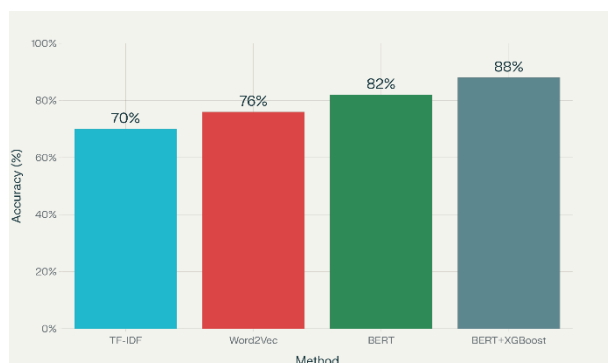


Fig. 1. Evolution of CO-PO Mapping Systems (Source: Adapted from the authors' analysis of the literature)

A. Manual and Rule-Based Systems

The foundational approaches were entirely manual, dependent on faculty expertise and predefined rubrics. While simple to implement, these methods were inherently subjective and non-scalable. Subsequent **rule-based systems** introduced automation through keyword matching and dictionary lookups. For instance, Kuruvila et al. proposed a system using Bloom's taxonomy for verb-level matching and WordNet for semantic analysis. Although this reduced manual effort, it lacked the linguistic intelligence to handle paraphrasing and complex academic language effectively.

B. Machine Learning-Based Systems

The second generation employed classical ML and statistical NLP. Gupta and Mehta [4] used TF-IDF vectorization combined with cosine similarity, quantifying similarity but relying on surface-level lexical patterns. Advancements introduced word embedding models like **Word2Vec** and Doc2Vec. Jain et al. [5] demonstrated a framework where COs and POs were converted into average Word2Vec vectors. While an improvement, these embeddings remained context-agnostic, struggling with polysemy and the specific context of educational outcomes.

C. Deep Learning and Transformer-Based Systems

The current state-of-the-art leverages deep learning, particularly transformer architectures. Zaki et al. [3] achieved 83.1% accuracy in automating CLO-PLO mapping using a Bag-of-Words model with cosine similarity, notably outperforming BERT and SpaCy in their specific setup, which they attributed to their optimized thresholding and the limited, structured nature of their dataset. However, transformer models like **BERT** fundamentally change the game by generating dynamic, context-aware embeddings. Singh et al. proposed a BERT-based model that captures the bidirectional context of words, enabling it to understand that "analyze a system" and "evaluate a process" are semantically related even without keyword overlap. Our proposed framework builds upon this promise of transformers, aiming to overcome the limitations of previous ML approaches by leveraging deep semantic understanding and complementing it with a robust classifier.

III. The Proposed Framework

The proposed framework is designed as a modular, end-to-end system for automating OBE compliance. The overall architecture is depicted in Fig. 2.

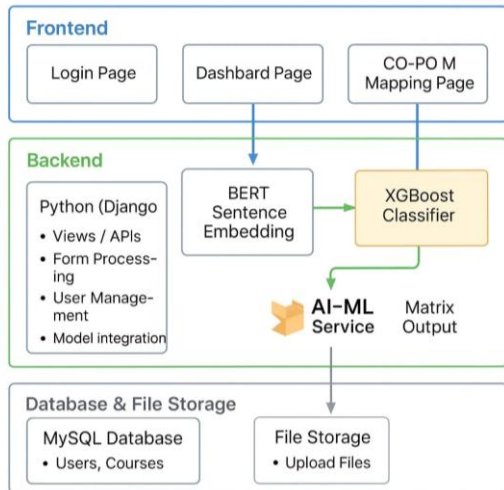


Fig. 2. System Architecture of the Proposed Framework
 (Source: Designed by the authors)

A. Input Module

The system accepts two primary inputs:

1. **Program Outcomes (POs):** list of m broad outcome\ statements\ for\ the\ program.
2. **Course Outcomes (COs):** collection of n courses, each with j_i specific outcome statements.

These are structured into a standardized format for processing, creating individual "(PO + CLOs)" tables for each combination [3].

B. Text Preprocessing Module

Raw text is cleaned and normalized through a standard NLP pipeline:

- **Tokenization:** Splitting sentences into individual words/tokens.
- **Lowercasing & Punctuation Removal:** Ensuring uniformity.
- **Stop Word Removal:** Filtering out common, non-informative words (e.g., "the", "a", "in").
- **Lemmatization:** Reducing words to their base or dictionary form (e.g., "analyzing" → "analyze") using libraries like SpaCy.

C. BERT-Based Semantic Feature Generation

This step extracts the deep semantic features necessary for accurate correlation.

1. **Embedding Generation:** Each pre-processed CO and PO statement is fed into a pre-trained **BERT model**. BERT generates a dense, contextualized embedding vector (typically 76 dimensions for **bert-base-uncased**) for the entire sequence, effectively capturing its semantic meaning [6]. These vectors serve as the rich feature set for the subsequent classification.
2. **Similarity Metric:** As an initial measurement, the semantic similarity between a CO embedding vector and a PO embedding vector is often quantified using Cosine Similarity. The formula is given by:

$$Similarity = \cos(\theta) = \frac{A \cdot B}{|A||B|}$$

where A and B are the BERT embedding vectors for the CO and PO, respectively. This results in a score between 0 and 1, providing a robust, context-aware metric of semantic closeness.

D. The eXtreme Gradient Boosting (XGBoost) Classification Layer

The similarity scores and, more powerfully, the raw **BERT embedding vectors** (of the CO and PO) are then combined and passed to the XGBoost module for the final classification of the correlation level (e.g., 0, 1, 2, or 3).

1. **XGBoost Principles:** eXtreme Gradient Boosting (**XGBoost**) is a highly optimized and scalable implementation of the gradient-boosted decision tree algorithm, operating as an effective **ensemble method**. It sequentially builds decision trees, with each new tree trained to **minimize the errors (residuals)** of the preceding models, iteratively refining the prediction capability. This process is formalized using a **second-order Taylor approximation** of the loss function, leading to faster convergence [3,4].
2. **Role in Mapping:** In this framework, XGBoost functions as the high-accuracy classification layer. It utilizes the **dense, contextualized vectors** from BERT as highly informative, pre-engineered features. XGBoost's ability to handle high-dimensional feature spaces and its intrinsic **regularization terms** (L1 and L2) are key to preventing overfitting to the training data, ensuring the final classification of the CO-PO alignment level is both accurate and stable [3,2, 3,4].

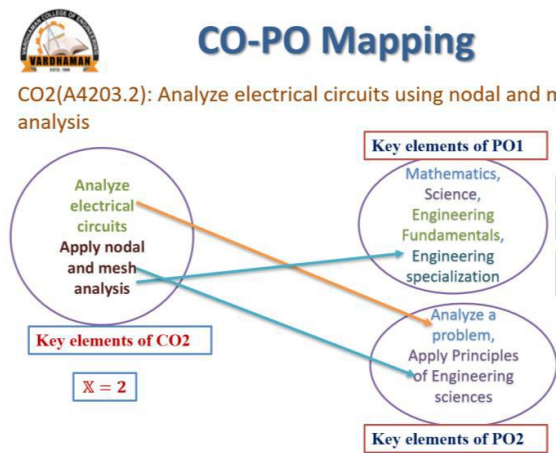
3. **Matrix Generation:** The XGBoost classifier predicts the final correlation level for every CO-PO pair. This process generates the final CO-PO correlation matrix, where each cell indicates the strength of the relationship.

E. Why BERT + XGBoost is Superior

The superiority of this approach stems from the synergy between the two models: **BERT** provides deep, **contextual semantic understanding** that defeats the limitations of keyword- and static-embedding models, and **XGBoost** provides a fast, robust, and regularized classification layer that accurately maps these complex semantic features into discrete correlation levels (0, 1, 2, 3), a task at which traditional classifiers often struggle [4.3, 4.5].

IV. Attainment Analysis Module

Mapping alone is insufficient for OBE; it must be linked to attainment measurement. The framework incorporates a detailed attainment analysis module, as shown in Fig. 3.



from [3])

A. Direct Attainment of Course Outcomes (COs)

CO attainment is calculated based on student performance in assessments directly linked to each CO.

- **Formative Assessment (CIE - 30%):** Includes Continuous Assessment Tests (CATs), assignments, and quizzes.
- **Summative Assessment (SEE - 70%):** The semester-end examination.

For each CO, the percentage of students scoring above a set academic threshold (e.g., 60% of the marks) is calculated. This direct attainment (DA_{CO}) is computed separately for CIE and SEE and then combined with their respective weightages.

B. Direct Attainment of Program Outcomes (POs)

PO attainment is derived from the CO attainment and the CO-PO correlation matrix generated by the combined BERT and XGBoost module.

$$Direct\ PO\ Attainment = \frac{\sum (CO\ Attainment \times Correlation\ Level)}{\sum Correlation\ Levels}$$

This is calculated for all courses contributing to a specific PO, providing a program-level performance metric.

C. Indirect Attainment

Indirect assessment provides a qualitative complement to direct metrics.

- **For COs:** Course End Survey (CES) where students self-report their perceived achievement.
- **For POs:** Student Exit Survey (SES), alumni surveys, and employer feedback.

The final PO attainment is often a weighted sum of direct (e.g., 90%) and indirect (e.g., 10%) components, offering a holistic view of program effectiveness.

V. Comparative Analysis and Discussion

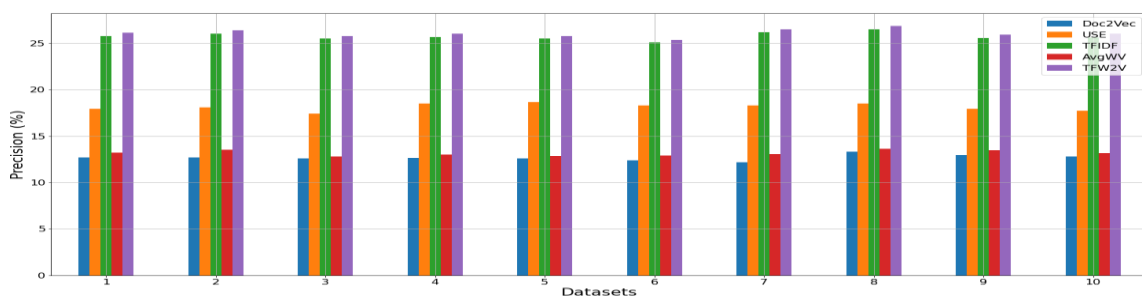


Fig. 4. Precision accuracy (Source: Adapted from [5])

A comparative analysis of CO-PO mapping techniques reveals a clear performance hierarchy driven by their underlying methodology. Early keyword-based systems like TF-IDF are limited to lexical matching, failing to capture semantic relationships and leading to inaccurate mappings where synonyms or paraphrases are involved.

Static word embedding models like Word2Vec improved upon this by understanding some semantic meaning, but their context-agnostic nature remained a critical flaw. They could not disambiguate words with multiple meanings, which often resulted in incorrect correlations.

The adoption of transformer-based models like BERT marked a significant advancement. By generating dynamic, context-aware embeddings, BERT can grasp the nuanced intent behind educational outcomes, accurately identifying semantic similarities even without keyword overlap. The proposed framework enhances this further by using BERT's sophisticated embeddings as input for an XGBoost classifier. This synergy leverages deep semantic understanding for feature generation and a robust, regularized model for final classification, yielding superior accuracy and stability.

Furthermore, the integration of this high-quality mapping with the structured attainment analysis module creates a **closed-loop system** essential for Continuous Quality Improvement (CQI), enabling institutions to identify gaps and implement targeted interventions.

VI. Future Work

The survey of existing literature reveals a rapidly evolving field with several promising avenues for future research. The next significant stride will be the incorporation of Explainable AI (XAI) into mapping frameworks. While models like BERT and XGBoost offer high accuracy, their decision-making processes are often opaque. Developing mechanisms to visually highlight the key semantic features or words that led to a specific CO-PO alignment is crucial. This transparency will be instrumental in building trust among educators and accreditation bodies, moving automation from a black-box tool to an interpretable assistant.

Another compelling direction is the move towards domain-specific fine-tuning of large language models. The current use of general-purpose models like BERT is effective, but their performance can be further enhanced by training them on a large, specialized corpus of educational outcome statements, syllabi, and accreditation manuals from engineering and other disciplines. This would allow the models to grasp the

nuanced terminology and specific contextual meanings within academic settings, leading to even more precise and reliable mappings.

Finally, to bridge the gap between research and practical application, the development of deployable and accessible tools is essential. Future work should focus on creating user-friendly web applications that allow educators to leverage these advanced AI models without any programming knowledge. Concurrently, research into optimizing these computationally intensive models for efficiency is vital.

VII. Conclusion

This comprehensive survey has charted the evolution of CO-PO mapping methodologies in Outcome-Based Education, from manual and rule-based systems to the current state-of-the-art leveraging deep learning and transformer models. The journey highlights a clear trajectory towards increasing automation, objectivity, and semantic understanding. Early keyword-matching techniques, while pioneering, proved limited in handling the complexity and nuance of educational outcomes. The advent of static word embeddings marked an improvement, yet the contextual revolution brought by models like BERT has fundamentally elevated the potential for accurate and meaningful automation.

The integration of these sophisticated Natural Language Processing models with powerful machine learning classifiers, such as XGBoost, represents the current frontier. This synergy effectively addresses the core limitations of earlier systems by capturing deep semantic relationships and translating them into reliable correlation matrices. Furthermore, the extension of these mapping techniques into integrated attainment analysis modules creates a powerful, data-driven foundation for continuous curriculum improvement, directly supporting the strategic goals of educational institutions and their accreditation processes.

In summary, the research domain of AI-driven OBE mapping is poised for significant impact. The move from purely theoretical models to practical, explainable, and efficient systems will define its future success. The collective work surveyed in this paper marks a definitive shift away from subjective, labour-intensive manual processes towards a future of intelligent, reliable, and data-informed quality assurance in higher education, empowering educators to make more strategic and effective decisions for curriculum enhancement.

References

- [1] N. Zaki et al., "Automating the Mapping of Course Learning Outcomes to Program Learning Outcomes using Natural Language Processing," *Research Square*, 2022.
- [2] J. S. Kuruvila et al., "An Automated Mapping Process Using NLP Technique To Correlate Program Objectives And Outcomes," *International Journal of Advanced Research*, vol. 4, no. 8, 2016.
- [3] B. R. Reddy et al., "Case Study on the Assessment of Program Quality through CO-PO Mapping and its Attainment," *Journal of Engineering Education Transformations*, vol. 34, 2021.
- [4] Mr. Ram Sarnaik et al., "REVIEW PAPER ON: CO-PO MAPPING AN ATTAINMENT FOR COURSES OF UNIVERSITY AFFILIATE ENGINEERING PROGRAMS," *International Journal of Progressive Research in Engineering Management and Science*, vol. 04, Issue 11, 2024.
- [5] Q. Duong et al., "TFW2V: An Enhanced Document Similarity Method for the Morphologically Rich Finnish Language," *Proceedings of the NLP4DH Workshop*, 2021.
- [6] A. Alshangiti et al., "Rule-based Approach toward Automating the Assessments of Academic Curriculum Mapping," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 12, 2020.